Exhibit AA

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of: SALK, Jesse

Application No.: 16/411,066

Filed: May 13, 2019

For: METHODS OF LOWERING THE ERROR RATE OF MASSIVELY PARALLEL DNA SEQUENCING USING DUPLEX CONSENSUS SEQUENCING

Confirmation No.: 2545

Art Unit: 1636

Examiner:

SECOND PRELIMINARY AMENDMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Commissioner:

Prior to examination on the merits, please amend the application as follows:

Amendments to the claims begin on page 2 of this paper.

Remarks begin on page 8 of this paper.

AMENDMENTS TO THE CLAIMS

The following complete listing of claims replaces all previous listings of claims in the application.

1-36. (Canceled)

37. (New) A method, comprising:

a) providing a population of circulating DNA molecules obtained from a bodily sample from a subject;

b) converting the population of circulating DNA molecules into a population of non-uniquely tagged parent polynucleotides, wherein each of the non-uniquely tagged parent polynucleotides comprises (i) a sequence from a circulating DNA molecule of the population of circulating DNA molecules, and (ii) an identifier sequence comprising one or more polynucleotide barcodes, such that each non-uniquely tagged parent polynucleotide is substantially unique with respect to other non-uniquely tagged parent polynucleotides in the population;

c) amplifying the population of non-uniquely tagged parent polynucleotides to produce a corresponding population of amplified progeny polynucleotides;

d) sequencing the population of amplified progeny polynucleotides to produce a set of sequence reads;

e) mapping sequence reads of the set of sequence reads to a reference sequence;

f) grouping the sequence reads into families, each of the families comprising sequence reads comprising the same identifier sequence and having the same start and stop positions, whereby each of the families comprises sequence reads amplified from the same non-uniquely tagged parent polynucleotide;

g) at each genetic locus of a plurality of genetic loci in the reference sequence, collapsing sequence reads in each family to yield a base call for each family at the genetic locus; and

h) determining a frequency of one or more bases called at the locus from among the families.

38. (New) The method of claim 37, further comprising detecting, at one or more loci, one or more of at least one single nucleotide variant and at least one copy number variant.

39. (New) The method of claim 37, wherein converting comprises any of blunt-end ligation, sticky end ligation, PCR, ligation-based PCR, single strand ligation and circularization to a single strand.

40. (New) The method of claim 37, wherein the reference sequence comprises a sequence from a human reference genome.

41. (New) The method of claim 37, wherein determining the frequency comprises detecting a rare mutation.

42. (New) The method of claim 37, further comprising generating a set of consensus sequences from the sequence reads, wherein determining the frequency of one or more bases comprises detecting a presence of sequence variations in the set of consensus sequences compared with the reference sequence.

43. (New) The method of claim 37, further comprising filtering out sequence reads that fail to meet a quality threshold.

44. (New) The method of claim 37, further comprising selectively enriching regions from a genome or transcriptome of the subject prior to sequencing.

45. (New) The method of claim 37, further comprising removing a subset of the sequence reads from further analysis prior to (e).

46. (New) The method of claim 37, wherein the plurality of circulating DNA molecules includes a genetic variant having a variant frequency lower than about 1% or lower than about 0.01%.

47. (New) The method of claim 37, wherein the plurality of circulating DNA molecules includes a genetic variant having a variant frequency as low as about 0.01% or as low as about 0.03%.

48. (New) The method of claim 37, wherein an error rate of the base call of each family determined in step (g) is lower than about $1x10^{-6}$ or is as low as about $1.2x10^{-9}$.

49. (New) The method of claim 37, wherein an error rate of the base call of each family determined in step (g) is no more than about $1.5x10^{-4}$ or about $3.5x10^{-5}$.

50. (New) The method of claim 37, wherein the circulating DNA molecules are nucleic acid-based serum biomarkers.

51. (New) The method of claim 37, wherein the identifier sequence comprises a polynucleotide barcode selected from about 2 to about 256 distinct barcode sequences.

52. (New) The method of claim 37, wherein the identifier sequence comprises a polynucleotide barcode selected from more than 256 distinct barcode sequences.

53. (New) The method of claim 37, wherein the distinct barcode sequences are contained within a library generated from oligonucleotides comprising known sequences.

54. (New) The method of claim 37, wherein said non-uniquely tagged parent polynucleotide can be differentiated from other non-uniquely tagged parent polynucleotides using a combination of at least a first non-unique polynucleotide barcode at a first end of said circulating DNA molecule and a second non-unique polynucleotide barcode at a second end of said circulating DNA molecule.

55. (New) The method of claim 37, wherein the sequence reads are grouped into families based on i) the polynucleotide barcode and ii) at least one of: sequence information at a beginning of the sequence from the circulating DNA molecule and sequence information at an end of the sequence from the circulating DNA molecule.

56. (New) The method of claim 62, wherein the population of circulating DNA molecules comprises double-stranded molecule, and wherein the identifier sequence further comprises a strand identifier, and wherein, for each family of sequence reads amplified from the same non-uniquely tagged parent polynucleotide grouped in step (f), the method further comprises determining if the family has at least one sequence read from each strand of the double-stranded molecule using the strand identifier.

57. (New) The method of claim 37, wherein the circulating DNA molecules are double-stranded molecules, and wherein for each of a plurality of families, the method further comprises:

confirming the presence of at least one sequence read from each strand of the double-stranded molecule; and

comparing the at least one sequence read obtained from one strand to the at least one sequence read from the other strand to form a consensus sequence of the double-stranded molecule,

wherein the consensus sequence comprises only nucleotide bases at which the sequence of both strands of the double-stranded molecule are in agreement, such that a base call occurring at a particular position in the consensus sequence is

identified as a true base call.

58. (New) A method, comprising:

a) attaching a set of molecular tags to a population of circulating DNA molecules obtained from a bodily sample of a subject to produce a population of tagged original DNA molecules, wherein a plurality of the tagged original DNA molecules has identical molecular tags, and wherein each tagged original DNA molecule is substantially unique with respect to other tagged original DNA molecules in the population;

b) amplifying the population of tagged original DNA molecules to produce a corresponding population of DNA molecule amplicons;

c) sequencing the population of DNA molecule amplicons to produce a set of sequence reads;

d) mapping sequence reads of the set of sequence reads to a reference sequence;

e) grouping the sequence reads into families based on i) the molecular tag and ii) sequence information derived from the circulating DNA molecule, whereby each of the families comprises sequence reads amplified from the same tagged original DNA molecule;

f) for each genetic locus in a set of genetic loci in the reference sequence, collapsing sequence reads in each family to provide an error-corrected consensus sequence read for each family at the genetic locus; and

g) determining a frequency of one or more genetic variants at the locus from among the families.

59. (New) The method of claim 58, further comprising selectively enriching regions from the subject's genome or transcriptome prior to sequencing.

60. (New) The method of claim 58, selectively enriching at least one of tagged original DNA molecules and DNA molecule amplicons for a subset of tagged polynucleotides that map to one or more genetic loci in the reference sequence.

61. The method of claim 58, further comprising detecting, at one or more loci, at least one single nucleotide variant or at least one copy number variant.

62. The method of claim 58, wherein determining the frequency comprises detecting a rare mutation.

63. (New) The method of claim 58, wherein the circulating DNA molecules are nucleic acid-based blood biomarkers.

64. (New) The method of claim 58, wherein the circulating DNA molecules are derived from neoplastic cells.

65. (New) The method of claim 58, wherein the set of molecular tags comprises about 2 to about 256 distinct molecular tags.

66. (New) The method of claim 58, wherein the set of molecular tags comprises more than 256 distinct molecular tags.

REMARKS

Claims 1-20 are cancelled in this amendment without any prejudice or disclaimer. Claims 21-36 were previously cancelled. Applicant reserves the right to pursue the subject matter of any canceled claims in the present application or in one or more continuing applications. Claims 37-66 are newly added with support throughout the original specification and claims. Because no new matter is introduced, entry and examination of the amended claims presented herewith are respectfully requested.

Dated: May 24, 2019

Respectfully submitted,

Customer No. 94991
Perkins Coie LLP
Patent - LA
P.O. Box 1247
Seattle, WA 98111-1247
Phone: (310) 788-3349
Fax: (206) 332-7198

PERKINS COIE LLP

By: /Lara J. Dueppen/
Lara J. Dueppen
Registration No. 65,002